European Research Studies Journal
Volume XXIII, Special Issue 1, 2020

pp. 660-678

Correlation of Economic Development of Countries with the Potential of their Housing Real Estate Markets: A Case Study in the European Union

Submitted 23/08/20, 1st revision 14/09/20, 2nd revision 25/10/20, accepted 11/11/20

Ireneusz Dąbrowski[1], Łukasz Mach[2], Łukasz Mikołajczyk[3], Arkadiusz Kuświk[4]

Abstract:

Purpose: *The aim of the research is to examine the correlation between economic development in the macroeconomic scale and the development of the real estate market.*
Design/Methodology/Approach: *The tool of the applied computational engineering is cluster analysis, the hierarchical method based on Euclidean distance and Ward's method (algorithm).*
Findings: *The research process carried out confirmed the existence of a dependence between macroeconomic determinants and parameters of the real estate market. A division into three homogeneous groups of countries, whose development shows signs of similarity, was obtained. This allows making comparisons between individual countries, neglecting their geographical and historical dependences.*
Practical Implications: *The utilitarian usability of applying computational tools used in economic analyses was proved.*
Originality/Value: *Individual clusters make a certain reference point for particular countries. The obtained results offer the possibility of applying a suitable strategy which provides bases for development and also the possibility of verifying its effects against a whole group.*

Keywords: *real estate market, economy, segmentation, cluster analysis, development, Ward's method, Euclidean distance*

JEL classification: *C38, R30.*

Paper Type: *Research study.*

Acknowledgement: *The paper presents personal opinions of the authors and does not necessarily reflect the official position of the Narodowy Bank Polski or the Warsaw School of Economics, or the Opole University of Technology, or the WSB University in Wrocław.*

[1]*Warsaw School of Economics, Collegium of Management and Finance, Poland, and the Narodowy Bank Polski, Research and Financial Innovation Department, Poland, e-mail: ireneusz.dabrowski@sgh.waw.pl;*
[2]*Opole University of Technology, Faculty of Economics Management, Poland, and the Narodowy Bank Polski, Regional Branch in Opole, Poland, e-mail: l.mach@po.edu.pl;*
[3]*Corresponding author, WSB University in Wrocław, Narodowy Bank Polski, Regional Branch in Opole, Poland, e-mail: lukasz.mikolajczyk@wsb.wroclaw.pl;*
[4]*Narodowy Bank Polski, Regional Branch in Opole, Poland, arkadiusz.kuswik@gmail.com;*

1. Introduction

The real estate market is – for many reasons – exceptional. The object of the turnover is an article which is specific, while one of its chief characteristics is its tight connection with the place (the ground) on which it is sited. The placement of a property very often constitutes one of the key elements which impact the interest on the part of the potential buyer (Trojanek, 2009). Bearing this in mind and with the help of a relevant research process, it was attempted to create homogeneous groups of European countries which are characterized by a similar rate of development. The hypothesis of a correlation existing between economic development and development of the property market was confirmed, as well. Establishing homogeneous clusters offers the possibility of comparing countries which develop in a similar way as far as the variables accepted in the research process are concerned, with one another, omitting in this way their geographical location or historical conditions. This, in turn, offers the possibility of accepting an optimal development strategy for countries that already enjoy a high level of development and require accepting new conceptions with reference to those which try to catch up with the better-developed states. An additional asset here is the chance to compare countries with those which are actually like them.

The development of the real estate market makes one of the main factors of economic development, hence influencing the development of this market allows raising justifiable expectations connected with economic development as such. Even though in the property market there occur similar dependences to those found in any other one, the former is rather not transparent. This poses a fair amount of difficulty to all kinds of analyses which are carried out with reference to this market. Exemplary analyses of the real estate market can be found, among others, in Cartone *et al.* (2020), where the determinants of economic development for 187 regions in 12 European countries were analyzed. Apart from other things, it was pointed out that the differences between EU regions, regarding economy, demand running an ad hoc policy. In turn, in the research reported in the work by Jílková and Skaličková (2019), the main goal is to indicate areas which should be taken into account while assessing economic development and affluence, in particular – social and environmental factors. The authors also pointed to methods of measuring econmic development, which are an alternative in relation to the most frequently used economic indexes that measure economic development. A more local analysis of economic developoment can be found in the work by Cudo (2019), where an attempt was made at determining the level of economic development of the smallest cities in West-Pomeranian Province (Poland) on the basis of selected indexes over time.

The property market very often possesses specific features which concern a certain narrower group connected with the division of this market, according to the spatial criterion. Analyses which illustrate taking this criterion into account can be found, among others, in the work of Brzezicka *et al.* (2018), where the situation in the

Polish housing real estate market as a developing market is analyzed, with special account taken of the market demand and supply. In turn, Turpenaite *et al.* (2017) distinguish the most important determinants of the fluctuations in the housing market in Lithuania in the years 2005-2015. An analysis of the cycle in the housing market is carried out and changes in prices and main factors impacting the changes are discussed. Then, the analysis conducted by Antoniucci and Marella (2018) allowed making an evaluation of changes in the polarization of the housing market in Italian cities beginning with 2008, that is in the years of economic recession in this country.

Linking the property market with economy is most justified. The former can exert an influence on the latter, even if through its share in the GDP and improvement in the quality of housing resources by commissioning new investments or its share in local taxes, as well. In return, economy also impacts the real estate market in that the height of interest rates is of significance to borrowers, both on the demand and the supply sides, a general boom in economy translates into an increase in the demand for property, incomes grow and this, in turn, means a growing sense of security and stability.

An economic situation does not develop unchangingly as it is subject to certain fluctuations. It is determined by economic cycles, which is treated about to a broader extent by Huang *et al.* (2020) or Chang (2019) in their works, where housing property markets are analyzed, taking into consideration cycles that occur in them. In turn, in the work by Łaszek and Olszewski (2018), the real estate market in Warsaw is analyzed and the results allow the authors to conclude that the property market is of the local character and, therefore, regional factors exert a very strong influence on its evolution.

Nonetheless, the global factors also play a vital role in this respect and a rapid economic growth can lead to strong cycles in the market and cause problems with stability of economy. Periods of economic development which also relates to the real estate market, as well as periods of recession or even those of crisis, are equally frequently subject to analyses. In the work by Mach (2019), the author undertakes to measure and assess the impact of a global crisis on the property market. As a result, a measurement of effects and intensity of their impact on the number of building permissions issued, the number of the employed in the construction industry, value of the building production and prices of real property was carried out. In turn, Li and Wei (2020), in their work, analyze local geographical areas, where there followed a drop in the value of housing (2008-2012) and a period of boom (2012-2016), beginning with the time of financial crisis, laying stress clearly on their impact on urban inequalities in Salt Lake County in the State of Utah.

In the literature on the subject, it is possible to find a great deal of research whose aim is to form homogeneous groups. For example, the studies of Zarikas *et al.* (2020) deal with the very current problem of COVID-19 pandemic. The

consequence of the research process is establishing a group of countries with reference to active cases, active cases on population and active cases on population and area based on Johns Hopkins epidemiologic data. Another instance of applying cluster analysis is the work by Calvo-Porral *et al.* (2020). It follows from the analysis of clusters based on data from 1,269 consumers of wine that there does not exist an average wine consumer. It is possible to distinguish four clusters dependent of emotions which accompany consumption of wine. It is worth paying attention to co-sharing economy, too. In the work of Lutz and Newlands (2018), the authors make a segmentation of consumers within one platform of co-sharing economy, Airbnb. Two different types of accommodation offered on Airbnb are compared a shared room and a whole house/flat.

Concluding, the use of analyses leading to segmentation finds a broad application in various domains. For the needs of realizing research goals, in the first place, the process of selection of diagnostic variables is carried out, then the accepted research assumptions are presented, on the basis of which the whole process of conducting relevant studies is accomplished. The summary of the studies is contained in the synthetic conclusion presenting the most significant findings and indicating their utilitarian application.

While reviewing the literature, one can also come across studies which aim to form homogeneous groups founded on data relating to the real estate market. In the work by Mikołajczyk (2017), the research process led to establishing two groups based on data concerning a division of the country into provinces (voivodeships). A slightly modified approach to segmentation can be found in the work of Wu *et al.* (2018) dealing with the housing market in Shenzhen in China. Studies in which the analysis of clusters was used as one of the steps to achieve the aim, were presented in the work of Calka (2019). The research process at the first stage distinguishes local property markets in which there are found those that are the most like one another as regards selected structural features.

2. Selection of Diagnostic Variables

The data used in the research process were obtained from Eurostat database. Because access to a part of them, chiefly those relating to the real estate market, is hampered or indeed impossible, it was decided to apply a two-stage approach. At the first step, variables were collected which characterize states as regards macroeconomics. On their basis there was carried out a process of grouping for the year 2015. The next step was repeating this process, yet with the use of data for the year 2018 which is the last one for which the data are available with reference to most countries. The second stage of analyzing the countries was executed based on dynamics. Macroeconomic variables on which the division was made into homogeneous groups was recalculated to dynamics covering three years. Because of the obtained values there were established groups which are characterized by a similar rate of development as regards the analyzed variables. The obtained results

were compared with dynamics relating to the data coming from the property market of each of the states. This allowed determining whether the property market existing in the groups formed earlier as regards macroeconomy is developing in an analogous manner. Or maybe it is necessary to make a separate grouping for the data that come from the real estate market?

It was decided that in the research process the following were important: the availability of data for the greatest possible number of countries and the number of features which characterize the states selected for the analysis. At the initial stage of the research process the following countries were excluded from the analysis: Greece – because of the lack of data concerning changes in prices in the property market, Montenegro, North Makedonia, Albania, Serbia and Turkey – since in the case of these states there was a shortage of data with a larger number of variables, hence the decision to eliminate them. In the case of Cyprus, it was only the value of surplus for the year 2015 which was missing; therefore, it was decided to complement the shortage and calculate it as a mean of two neighboring values. Norway also provided one variable which did not have the value of the GDP for the year 2018. Still, it was finally obtained from another source and successfully complemented. Eventually, there were 28 countries which were taken for the analysis.

In search of the value of potential variables there was made a review of Eurostat, EBC and OCDB database. The following initial diagnostic variables which were meant to be used were obtained from Eurostat database:

1. Gross domestic product (GDP) at current market prices by NUTS 3 regions [nama_10r_3gdp],
2. Unemployment rates by sex, age and NUTS 2 regions (%) [lfst_r_lfu3rt] – (20-64 years),
3. Live births (total) by month [demo_fmonth],
4. Deaths (total) by month [demo_mmonth],
5. Population on 1 January by age and sex [demo_pjan],
6. Population: Structure indicators [demo_pjanind] - median age of population,
7. Annual net earnings [earn_nt_net] - Single person without children earning 100% of the average earning,
8. HICP (2015 = 100) - annual data (average index and rate of change) [prc_hicp_aind],
9. Employment and activity by sex and age - annual data [lfsi_emp_a] – (20-64 years).

The above list was subjected to initial analysis which was meant to exclude all the data of little variability. The assumption was accepted that the value of the variability index below 10% excludes the given variable from further analysis. Ultimately, only variable 6 was excluded from the set. Another step was checking

whether there exists a strong correlation between some of the variables. Thus, it was assumed that the value of a correlation coefficient exceeding 0.7 causes elimination of one of the variables which the correlation concerns. On conducting the analysis of all coefficients, the following variables were removed from the set: 3, 4, 5 and 9. Thus, the final list of diagnostic variables presents itself as follows:

1. Gross domestic product (GDP) at current market prices by NUTS 3 regions [nama_10r_3gdp],
2. Unemployment rates by sex, age and NUTS 2 regions (%) [lfst_r_lfu3rt] – (20-64 years),
3. Annual net earnings [earn_nt_net] - Single person without children earning 100% of the average earning,
4. HICP (2015 = 100) - annual data (average index and rate of change) [prc_hicp_aind].

3. Methodology of the Research

The variables expressed in different units should be normalized so that they could be comparable. There exist several formulas which can be used; however, in consequence of the comparative research results relating to normalization (Milligan and Cooper 1988) it was decided to apply zero unitarization:

$$x'_i = \frac{x_i - min.x_i}{max.x_i - min.x_i} \qquad \text{for i = 1,…,n} \tag{1}$$

The results which were attained in the above-mentioned research included eight strategies of normalization and the one pointed to above turned out to be the best. Because the main aim of the research process is to obtain the most homogeneous groups, it was not established at the beginning how many such groups should be formed. The number was to be fixed based on selected indexes and a tree diagram that was to be formed based on the agglomerative grouping method. The indexes selected to establish the optimal number of clusters are the following (Walesiak and Gatnar 2009):

1. Silhouette

$$S(u) = \frac{1}{n}\sum_{i=1}^{n} \frac{b(i)-a(i)}{\max\{a(i);b(i)\}} \; ; S(u)\epsilon[-1;1] \tag{2}$$

where $a(i) = \frac{\sum_{j\epsilon C_k; j\neq i} d(i,j)}{(n_k-1)}$ can be defined as the mean distance of object i from all the other objects in cluster C_k (thus the one which includes object i); element $b(i) = \min_{C_r \neq C_k} d(i,C)$, where $d(i,C) = \frac{\sum_{q\epsilon C_r} d(i,q)}{n_r}$ is the minimal mean distance of object i belonging to cluster C_k from the remaining ones.

The maximal value of index $arg \max_{u}\{S(u)\}$ should be chosen.

2. Hartigan's

$$H(u) = \left(\frac{W_u}{W_{u+1}} - 1\right)(n - u - 1)\ ;\ H(u)\epsilon R_+ \tag{3}$$

where W_u – intra-class covariance matrix, u – number of classes, n – number of objects. While deciding about the choice of a concrete number of classes, the lowest u for which $H(u) \leq 10$ should be chosen.

3. Gap

$$Gap(u) = \frac{1}{B}\sum_{b=1}^{B} logW_{ub} - logW_u\ \ ;Gap(u)\epsilon R \tag{4}$$

$$diffu = Gap(u) - Gap(u+1) + s_{u+1} \tag{5}$$

$$s_u = sd_u\sqrt{1 + 1/B} \tag{6}$$

where B – number of generated sets of observations, Wu – intra-class covariance matrix, sd_u –standard deviation from the value log W_{ub}. The choice of the number of classes, which is optimal due to the Gap coefficient, should be based on the lowest value of u for which $diffu \geq 0$.

4. Krzanowski and Lai's

$$KL(u) = \left|\frac{DIFF_u}{DIFF_{u+1}}\right| \tag{7}$$

$$DIFF_u = (u-1)^{2/m}W_{u-1} - u^{2/m}W_u \tag{8}$$

where m – number of variables, u – number of classes, W_u – intra-class covariance matrix. The selection of the number of classes is made on the basis of the formula: $arg \max_{u}\{KL(u)\}$.

In order to determine the place of dividing the dendrogram it is possible to make use of a few methods. It was decided to apply Mojena's rule (more can be found in the work by Mojena (1977)). It allows marking out the point of "cutting off" on the basis of bonds by means of the following formula:

$$d_{i+1} > \bar{d} + ks_d \tag{9}$$

where:
d_i – are the distances between individual bonds,

$\bar{d}$ – mean from d_i,
s_d – standard deviation from d_i,
k – constant, the author of the rule (Mojena) suggests that it should be a value from the range between 2.75 to 3.50; however, in the work (Milligan & Cooper 1985) one can find the suggestion that the optimal value is 1.25.

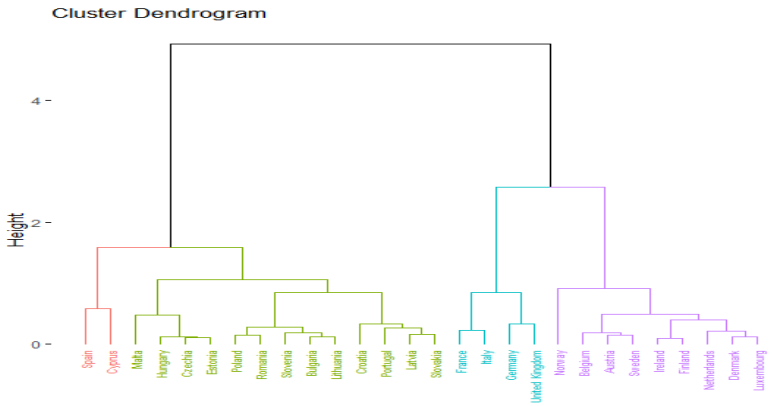
The number of indexes which can be used to establish the optimal number of groups is greater; nevertheless, in the accepted research process it was decided to apply the ones indicated above. The accepted schema of proceeding looks as follows:

Step 1 – selecting variables based on accepted assumptions relating to their variability and/or correlation.
Step 2 – selecting the optimal number of groups.
Step 3 – dividing countries into clusters according to the accepted measure of distance and the algorithm which makes the division (using a dendrogram).
Step 4 – describing statistically the obtained results.

4. Process of Research Realization

After normalizing the data for the year 2015, the values of indexes were calculated. To do so, program R was used, clusterSim packet, which was described in the work of Marek Walesiak and Dudek (2006) and in the primary source (Marek, Walesiak, and Dudek 2019). The obtained results indicate that the number of clusters should be established at four. The Euclidean distance was applied as it reflects this measure well for data expressed on the ratio scale, as well as Ward's method. The choice of the method was made based on the research conducted by Milligan and Cooper (1987), where the algorithm was most often pointed to as the best.

Figure 1. *Dendrogram for the data of 2015 together with the division into four clusters*



Source: *Authors' own elaboration with the use of program R.*

The agglomerative method allowed making a division which is presented in the following way:

Group 1: Spain, Cyprus,
Group 2: Malta, Hungary, Czechia, Estonia, Poland, Romania, Slovenia, Bulgaria, Lithuania, Croatia, Portugal, Latvia, Slovakia,
Group 3: France, Italy, Germany, United Kingdom,
Group 4: Norway, Belgium, Austria, Sweden, Ireland, Finland, Netherlands, Denmark, Luxemburg.

For the purpose of giving their brief characteristics, the basic statistics for all the four groups are shown in Table 1.

Table 1. *Basic statistical measures for the groups (data as of 2015)*

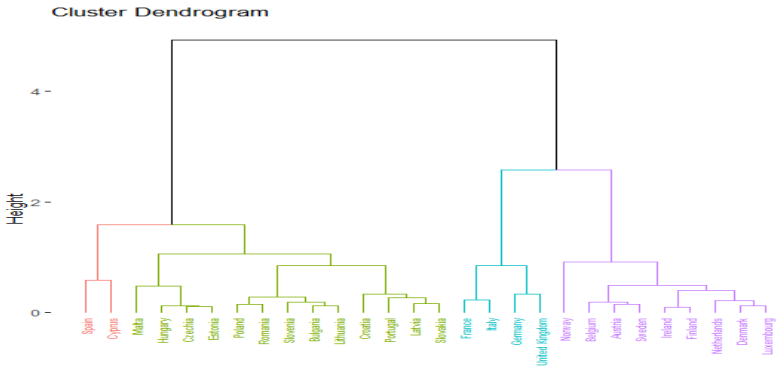
	Mean from the GDP [nama_10r_3gdp] mln Euro	Mean from unemployment rates % [lfst_r_lfu3rt] 20-64 years	Mean from annual net earnings [earn_nt_net]	Mean from HICP (2015 = 100) – [prc_hicp_aind]
Group 1	547708.475	18.300	20746.230	-1.050
Group 2	104003.425	8.715	8831.086	-0.146
Group 3	2381197.910	7.725	28080.880	0.225
Group 4	339136.417	6.867	33717.068	0.489

Source: *Authors' own calculations on the basis of Eurostat data*

Analyzing the individual groups with respect to all the variables, it can be observed that Groups 1 and 3 have the GDP above the mean from all the analyzed countries. The unemployment rate is lower than the mean of Group 3 and 4. The inflation level (HICP) below the mean is recorded in groups in which it is negative, that is Groups 1 and 2. However, it is worth paying attention to the fact that a low, yet still positive inflation is more desirable, hence Groups 3 and 4 point to a more suitable level for development of economy, although it is most frequently acknowledged that it should amount to about 2-3% (Nowak and Zalega, 2015). The annual net income obtained by one person without children, who earns 100% of the average net earnings was higher than the average in Groups 1, 3 and 4; however, in Group 1 it was higher only by 1.5%.
Taking into consideration the data from 2018, therefore valid for three years afterwards, the values of the indexes of the optimal number of clusters suggest a division into three clusters. The period of three years is not too long, though, which can suggest that there should not occur any greater changes. However, it can be seen that as early as at the beginning the optimal division points to a different number of groups. The groups are shown in Figure 2 which presents a dendrogram delineated on the basis of the same variables, yet for the values from the year 2018.
The division into three clusters, with the use of the agglomerative method, looks as follows:

Group 1: Estonia, Romania, Latvia, Lithuania, Slovakia, Bulgaria, Hungary, Poland, Czechia, Malta, Slovenia, Cyprus, Croatia and Portugal,
Group 2: Finland, Denmark, Ireland, Norway Belgium, Austria, Sweden, Luxemburg and Netherlands,
Group 3: Germany, United Kingdom, Spain, France and Italy.

Figure 2. *Dendrogram for the data from 2018, including the division into three clusters*



Source: *Authors' own elaboration with the use of program R*

The basic statistics relating to all the three groups are shown in Table 2.

Table 2. *Basic statistical measures for the groups (data as of the year 2018)*

	Mean from the GDP [nama_10r_3gdp] mln Euro	Mean from unemployment rates % [lfst_r_lfu3rt] 20-64 years	Mean from annual net earnings [earn_nt_net]	Mean from HICP (2015 = 100) – [prc_hicp_aind]
Group 1	115957.554	5.429	11170.795	2.214
Group 2	375395.309	5.022	35482.736	1.733
Group 3	2217762.118	8.220	26696.556	1.880
Group 4	115957.554	5.429	11170.795	2.214

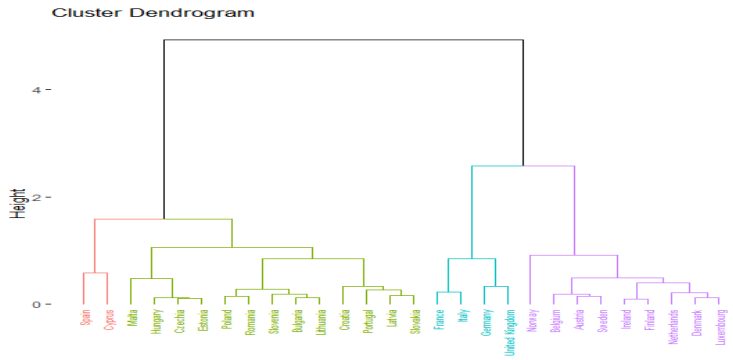
Source: *Authors' own calculations on the basis of Eurostat data.*

The two countries which formed a separate group in 2015 (Spain and Cyprus), this time were placed in two different groups. Analyzing the individual variables and obtained mean values in the groups (Table 2), it can be noticed that in the case of the GDP, Group 3 presents itself above the mean of all the countries. The unemployment rates which are lower than the mean one occurs in Groups 1 and 2 and remain close to each other. The inflation level (HICP) in all the groups is positive, Group 1 remaining above the general mean, while the values for the other groups are close to one another and only slightly below the mean. Group 1, again, is the closest to that of 3-4%, with the result amounting to 2.21%. The height of

annual net earnings that was obtained by one person earning the mean value, was above the mean for the whole set of countries in Groups 2 and 3, the highest level being recorded in the latter.

It can be noticed that Spain is the country which clearly lags behind the others (France and Italy) in Group 3, into which it was included as the last one. At the next step, this group was joined by another one, consisting of Germany and United Kingdom. On the other hand, Cyprus was included in the group composed of two countries: Croatia and Portugal, which prior to this had formed a smaller cluster and Cyprus found itself the nearest it. Analyzing the variables for Spain, the change of the GDP over those three years was smaller than it should result from the general mean, the unemployment rate dropped, although the decrease was lower than it would result from the general mean, the annual net earnings rose to a lesser degree than the mean from all the analyzed countries indicates, and – lastly – the inflation rate grew and its growth rate was similar to the general mean. In search of an answer to the question why in the year 2018 Spain and Cyprus – the countries which three years earlier had formed a separate cluster – found themselves in different groups, an attempt was made at grouping the countries due to dynamics of diagnostic variables. The process of data preparation was conducted earlier on the basis of absolute values, hence all the variables were left in the analysis based on dynamics. Despite the fact that it is expressed with the same unit, the decision was taken to normalize the values as they were characterized by a large stretch, with the minimal value amounting to 44% and the maximal one – 131%. Normalization of the data allows their standardizing and grouping due to the characteristics of the countries and not the weight of the variables. In order to determine the optimal number of clusters, the same indexes were applied, and the obtained results suggested making a division into three clusters. Accordingly, Figure 3 presents a dendrogram of such a division.

Figure 3. *Dendrogram for the values of dynamics, including the division into three clusters*



Source: *Authors' own elaboration with the use of program R.*

The above figure shows that Spain and Cyprus, which at first formed a separate cluster are characterized by a slightly different dynamics, since the two countries finally were qualified into different groups. This may suggest why this separation occurred in the year 2018. Thus, the division into three groups looks as follows:

Group 1: United Kingdom, Norway, Denmark, Finland, Sweden, France, Italy, Luxemburg, Austria, Belgium, Germany, and Spain,
Group 2: Netherlands, Portugal, Slovenia, Slovakia, Croatia, Poland, Malta, Ireland, and Cyprus,
Group 3: Estonia, Latvia, Lithuania, Bulgaria, Romania, Czechia, and Hungary.

The basic statistics relating to the dynamics of particular groups are shown in Table 3.

Table 3. *The basic statistical measures for the groups (dynamics)*

	Mean from the GDP [nama_10r_3gdp] mln Euro	Mean from unemployment rates % [lfst_r_lfu3rt] 20-64 years	Mean from annual net earnings [earn_nt_net]	Mean from HICP (2015 = 100) – [prc_hicp_aind]
Group 1	107.99	79.95	103.55	104.49
Group 2	117.44	56.30	109.24	102.56
Group 3	122.68	63.05	127.77	105.46
Group 4	107.99	79.95	103.55	104.49

Source: *Authors' own calculations on the basis of Eurostat data*.

Considering the obtained results, it can be stated that regarding the fact that the economic variables which were chosen for the analysis, as well as their dynamics during three years, it is evident that Europe does not form a monolith. We come to deal with a division of the countries into three groups, whose rates of development are varied. The analyzed period is too short to speak about stable divisions, though. This would demand conducting successive analyses over longer series of time spans. The mean dynamics of inflation in all the groups displayed a similar growth.

However, because this measure is neither a typical stimulant nor an inhibitor, it was decided that it should not be taken into account while establishing the ranking of the groups. The GDP and the annual net earnings were classified as stimulants, whereas the unemployment rate as an inhibitor. The position of each group was indicated for the above-mentioned three variables and the mean calculated from the positions points to that the fastest developing group is Cluster 3, Cluster 2 displays an average development rate and Cluster 1 is characterized by the slowest tempo of development. Such a division allows concluding that Spain, due to the fact that the country was characterized by the slowest development rate in comparison with Cyprus which was developing at a slightly faster pace, had to be placed in a

different group. Eventually, regarding the year 2018, while looking at the mean ranking relating to the analyzed variables, Spain was located among the states of a medium macroeconomic position, while Cyprus – to the group of the lowest one. Despite its higher development rate, Cyprus reached the level that made the country be assigned to the group of states characterized by the lowest economic standard. It is obvious that the country has a lot to make up for, yet the greater tempo suggests that Cyprus is getting nearer and nearer the more highly developed states.

Linking economy and its development with the situation existing in the real estate market is well-known in the literature (Kucharska-Stasiak, 2016), (Foryś and Mariańska, 2011). The all-economic activity is transferred onto the demand side of the property market. While a favorable economic situation lasts, also the supply side in the property market is more active. This can manifest itself through a rising number of new investments, higher prices, or standards of the housing area on offer but can also be seen in an alteration of expectations concerning the whole real estate market. If the economic boom includes an improvement in the situation of households, too, accomplished through a greater stability of employment, higher wages and a stronger sense of general stability, then this can lead directly to a wish of enhancing the housing conditions. The property market does not have to react to changes in economy at once, since this most often requires time.

Analyzing economic development of European countries over the span of three years, one can expect that the situation in the real estate markets of these countries should be characterized by similar tendencies. Accordingly, to verify this hypothesis, the authors decided to apply the division of the countries, which had been used before. Still in the statistics, the variables relating to the property market understood in a broader way, that is including construction industry and housing conditions, were considered. Upon verifying the availability of variables and their correlation, the following were established as the final ones:

1. House Price Index (HPI),
2. Building permits – number of dwellings,
3. Construction cost index,
4. Production in construction,
5. Overcrowding rate,
6. Housing cost overburden rate.

All the variables are expressed as dynamics and cover three years (2015-2018). The countries under analysis were divided into the same three groups as those established because of the analysis of the dynamics of macroeconomic variables. Because the absolute values were not available, it turned out necessary to base the correlation analysis on measures of dynamics. Thus, the housing cost overburden rate is a variable given in percent and means the proportion of population living in a household in which the total housing costs (after deducting housing benefits) constitute over 40% of the household's total available income. Then, the

overcrowding rate is a variable expressed in percent values. This index is defined as the percent of population living in an overcrowded household. It is considered that a given person lives in an overcrowded household if it does not have the minimal number of rooms at its disposal, equaling the following:

- one room per a household;
- one room shared between two persons in the household;
- one room for each single person at the age of 18 years and over;
- one room for a pair of persons of the same sex at the age spanning 12 to 17 years;
- one room for each person aged 12-17, not included in the category given above;
- one room for a pair of children under 12.

The mean values obtained for the variables relating to the real estate markets in individual clusters are presented in Table 4.

Table 4. *Mean values of variables relating to the property markets in the clusters*

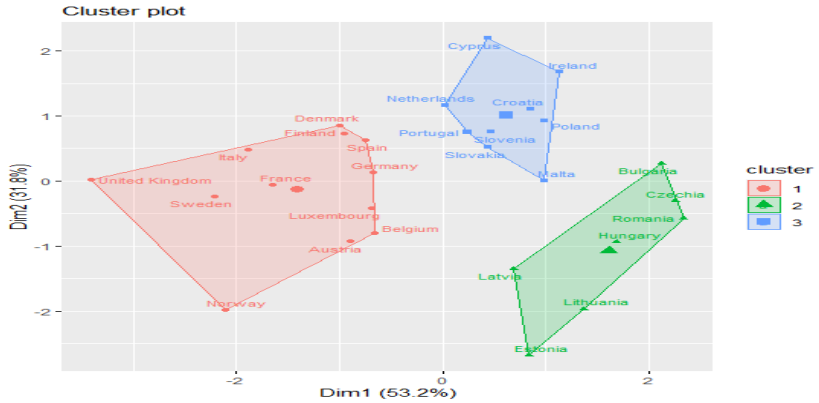
	Mean house price index (2015 = 100) annual data [prc_hpi_a]	Mean from building permits - number of dwellings	Mean from construction cost index - in national currency	Mean from construction production index - annual data [sts_copr_a]	Mean housing cost overburden rate by sex - EU-SILC survey [TESSI160]	Mean from overcrowding rate by sex - EU-SILC survey [TESSI170]
Group 1	113.17833	129.76666	105.966666	108.76666	102.60148	123.7320902
Group 2	119.27666	186.51111	104.566666	124.67777	74.597006	102.9186063
Group 3	126.80428	165.68571	110.771428	112.32857	82.392148	87.3898256

Source: *Authors' own calculations on the basis of Eurostat data*.

When considering which cluster is developing the fastest, each group was correlated with the GDP and then – on the basis of the sign of correlation coefficient – divided into stimulants and inhibitors. The sector of the property market and the very investments in the market themselves have their share in generating of the GDP. At the next step, for each variable there was made a ranking of groups and in the end the means were calculated in order to underline the order of individual clusters. The result confirmed that the countries of Group 3 developed the fastest, a medium pace characterized those in Group 2 and the lowest rate was typical of the countries of Group 1.

The division of all the countries into three clusters is presented in Figure 4. The individual groups and values of dynamics for the countries which belong to them were presented in Figures 5, 6 and 7.

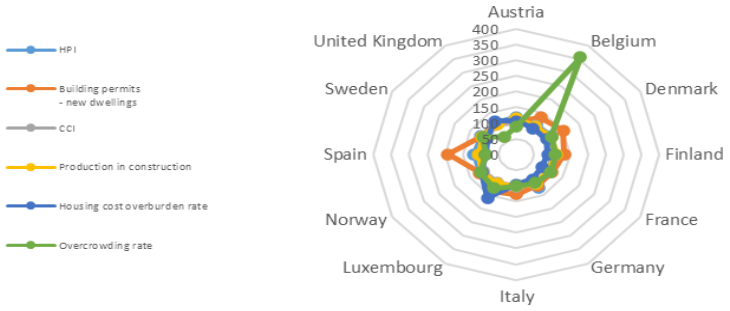
Figure 4. *Division of countries into three groups*



Source: *Authors' own elaboration on the basis of Eurostat data.*

Figure 4 shows the division of the countries into three groups due to the dynamics of their economic development, which was accomplished with the use of diagnostic variables.

Figure 5. *Countries and their dynamics – Group 1*



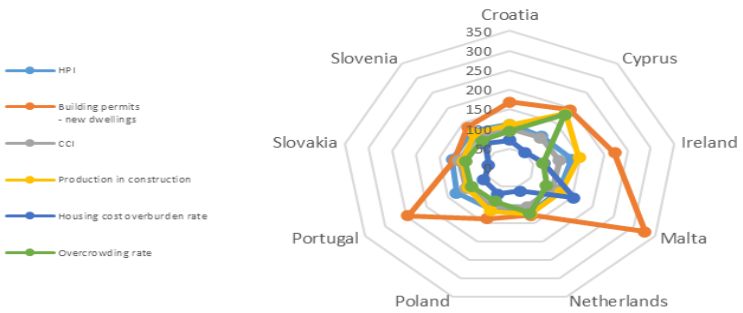
Source: *Authors' own elaboration on the basis of Eurostat data.*

While analyzing the countries composing Group 1 (Figure 5), it can be seen that they are similar to one another, with the only value concerning Belgium, i.e. the variable relating to overcrowding, which stands out. Although, in the case of this country, the value is not on a high level, the relevant percent rose by over three times during the analyzed three years.

Figures 6 and 7 show the countries included in Group 2 and Group 3, respectively. All the diagnostic variables relating to the real estate market point to a similar level of their dynamics, except the variable of 'building permits – new dwellings'. In both groups this variable is characterized by the greatest dynamics and visible differences in its values in particular countries. The variables concerning

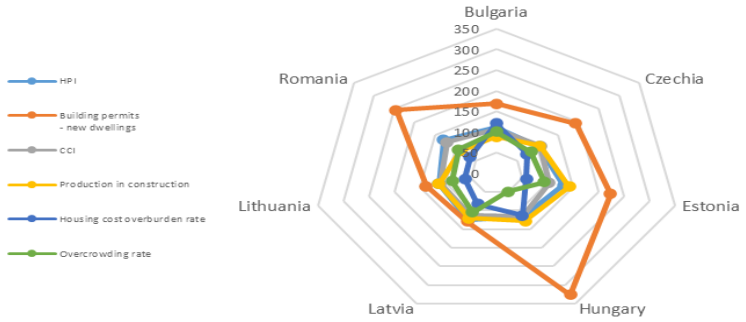
overcrowding and costs of housing, given as the percentage of households, also indicate high variability in individual countries.

Figure 6. *Countries and their dynamics – Group 2*



Source: *Authors' own elaboration on the basis of Eurostat data.*

Figure 7. *Countries and their dynamics – Group 3*



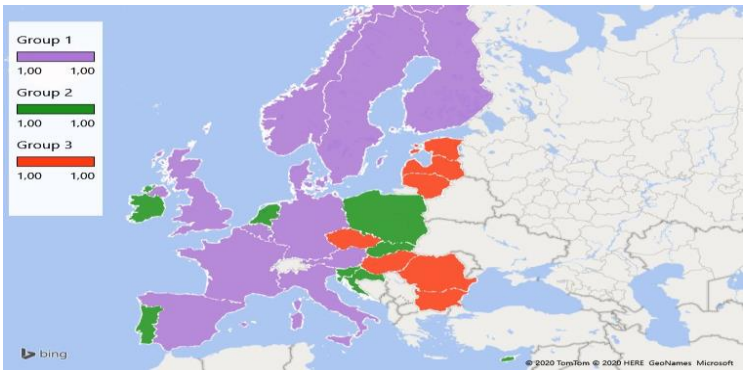
Source: *Authors' own elaboration on the basis of Eurostat data.*

Due to the fact that they have the character of inhibitors, this variability is expressed chiefly as a drop in the dynamics in the countries of Group 3; here, the medium index of dynamics shows a decrease. As regards Group 2, the majority of countries lowered the overcrowding index; however, as a result of a clear rise in two of the countries, it displays a slight growth. Looking at the next index which informs about housing costs, the countries of Group 2 recorded the greatest mean fall. On the other hand, the countries of Group 1 more often recorded an increase in the dynamics of these indexes. This may result from the fact that the property markets in these countries are developing a little slower than the economic development suggests. This, in turn, implies that the strategy of development of the real estate markets in these countries requires verifying or intensifying activities.

5. Conclusion

The conducted analysis focused on selecting homogeneous groups of European countries with reference to their development. The individual groups were presented regarding their geographical location in Figure 8. The results which were obtained in the analysis suggest that it is possible to take advantage of mutual bonds in economy.

Figure 8. *Division into three clusters shown with reference to geographical location*



Source: *Authors' own elaboration in Excel.*

A hypothesis was also presented, relating to the dependence between the development of the countries under analysis in the division into groups, their economic development, and the development of the broadly conceived real estate market. At the first stage, a division of the countries into three clusters was obtained, in this way confirming that the economic development of Europe comprises the values of individual entities which additionally form a cluster structure. In the case of data relating to the property market, building industry and housing conditions, a division into the same clusters was made and mean dynamics were analyzed to confirm the hypothesis.

Eventually, it was confirmed and the division of the countries into groups regarding macroeconomic variables which represent the economies is characterized by the same order of the groups concerning the rate of development. Poland belongs to Group 2, thus finds itself among the states showing an average pace of development. Information which can be drawn from segmentation of particular economies can be used to establish the right development strategy for individual groups or individual countries. Each state can compare itself to those in its own group and, in this way, observe whether the implemented actions aimed at raising the dynamics of its development bring intended effects. Still, it needs remembering that making comparisons with one's 'neighbors' may not offer the best solution as the geographical location provides no guarantee for similarities to exist between countries.

References:

Antoniucci, V., Marella, G. 2018. Is social polarization related to urban density? Evidence from the Italian housing market. Landscape and Urban Planning, 177, 340-349. https://doi.org/10.1016/j.landurbplan.2017.08.012.

Brzezicka, J., Wisniewski, R., Figurska, M. 2018. Disequilibrium in the real estate market: Evidence from Poland. Land Use Policy, 78, 515-531. https://doi.org/10.1016/j.landusepol.2018.06.013.

Calka, B. 2019. Estimating residential property values on the basis of clustering and geostatistics. Geosciences, 9(3), 1-14. https://doi.org/10.3390/geosciences9030143.

Calvo-Porral, C., Lévy-Mangin, J.P., Ruiz-Vega, A. 2020. An emotion-based typology of wine consumers. Food Quality and Preference, 79, 103777. https://doi.org/10.1016/j.foodqual.2019.103777.

Cartone, A., Postiglione, P., Hewings, G.J.D. 2020. Does economic convergence hold? A spatial quantile analysis on European regions. Economic Modelling. https://doi.org/10.1016/j.econmod.2020.03.008.

Chang, K. 2019. Are cyclical patterns of international housing markets interdependent? Economic Modelling. https://doi.org/10.1016/j.econmod.2019.09.002.

Cudo, M. 2019. The economic development of the smallest cities in West-Pomerania Province. Rozwój Regionalny i Polityka Regionalna, 45, 59-72. https://doi.org/10.14746/rrpr.2019.45.05.

Foryś, I., Mariańska, A. 2011. Socioeconomic determiants of the development of the housing market in Poland: a quantitative framework. Wydawnictwo Naukowe Uniwersytetu Szczecińskiego. https://books.google.pl/books?id=ljpvMwEACAAJ.

Huang, Y., Li, Q., Liow, K.H., Zhou, X. 2020. Is Housing the Business Cycle? A Multiresolution Analysis for OECD Countries. Journal of Housing Economics, 101692. https://doi.org/10.1016/j.jhe.2020.101692.

Jílková, E., Skaličková, J. 2019. Measurement of the Performance of an Economy. Zeszyty Naukowe Uniwersytetu Przyrodniczo-Humanistycznego w Siedlcach. Seria: Administracja i Zarządzanie, 47(47), 41-48. https://doi.org/10.34739/zn.2019.47.05.

Kucharska-Stasiak, E. 2016. The economic dimension of property. Wydawnictwo Naukowe PWN. https://books.google.pl/books?id=7ijFjwEACAAJ.

Łaszek, J., Olszewski, K. 2018. Regional Development of Residential and Commercial Real Estate in Poland and the Risk of Real Estate Cycles. Barometr Regionalny, 16(1), 41-51.

Li, H., Wei, Y.D. 2020. Spatial inequality of housing value changes since the financial crisis. Applied Geography, 115, 102141. https://doi.org/10.1016/j.apgeog.2019.102141.

Lutz, C., Newlands, G. 2018. Consumer segmentation within the sharing economy: The case of Airbnb. Journal of Business Research, 88, 187-196. https://doi.org/10.1016/j.jbusres.2018.03.019.

Mach, Ł. 2019. Measuring and assessing the impact of the global economic crisis on European real property market. Journal of Business Economics and Management, 20(6), 1189-1209. https://doi.org/10.3846/jbem.2019.11234.

Mikołajczyk, Ł. 2017. The division of the capital cities of provinces into homogeneous groups as regards their local real estate markets with the use of cluster analysis. Przegląd Nauk Stosowanych. Wybrane Problemy Rynku Nieruchomości, 14, 106-118.

Milligan, G.W., Cooper, M.C. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika, 50(2), 159-179.

https://doi.org/10.1007/BF02294245.

Milligan, G.W., Cooper, M.C. 1987. Methodology Review: Clustering Methods. Applied Psychological Measurement, 11(4), 329-354. https://doi.org/10.1177/014662168701100401.

Milligan, G.W., Cooper, M.C. 1988. A study of standardization of variables in cluster analysis. Journal of Classification, 5(2), 181-204. https://doi.org/10.1007/BF01897163.

Mojena, R. 1977. Hierarchical grouping methods and stopping rules: an evaluation. The Computer Journal, 20(4), 359-363. https://doi.org/10.1093/comjnl/20.4.359.

Nowak, A.Z., Zalega, T. 2015. Macroeconomy. Polskie Wydawnictwo Ekonomiczne S.A.

Trojanek, M. 2009. Purchasers prefernces in the primary housing market in Poznań. Acta Sci. Pol. Administratio Locorum, 8(1), 5-19. http://bazhum.muzhp.pl/media//files/Acta_Scientiarum_Polonorum_Administratio_Locorum/Acta_Scientiarum_Polonorum_Administratio_Locorum-r2009-t8-n1/Acta_Scientiarum_Polonorum_Administratio_Locorum-r2009-t8-n1-s5-19/Acta_Scientiarum_Polonorum_Administratio_L.

Tupenaite, L., Kanapeckiene, L., Naimaviciene, J. 2017. Determinants of Housing Market Fluctuations: Case Study of Lithuania. Procedia Engineering, 172, 1169-1175. https://doi.org/10.1016/j.proeng.2017.02.136.

Walesiak, M., Gatnar, E. 2009. Statistical analysis of data using program R. Wydawnictwo Naukowe PWN. https://books.google.pl/books?id=KWCoXwAACAAJ.

Walesiak, Marek, Dudek, A. 2006. Simulation-based optimalization of the choice of classification procedure for the given type of data - computer software and research results. Klasyfikacja i Analiza Danych - Teoria i Zastosowania, 13(1126), 120-129.

Walesiak, Marek, Dudek, A. 2019. Package 'clusterSim.' https://doi.org/10.1007/BF02294245>.

Wu, C., Ye, X., Ren, F., Du, Q. 2018. Modified data-driven framework for housing market segmentation. Journal of Urban Planning and Development, 144(4). https://doi.org/10.1061/(ASCE)UP.1943-5444.0000473.

Zarikas, V., Poulopoulos, S.G., Gareiou, Z., Zervas, E. 2020. Clustering analysis of countries using the COVID-19 cases dataset. Data in Brief, 31, 105787. https://doi.org/10.1016/j.dib.2020.105787.